基于双模型集成的太阳黑子磁类型分类

陈清源[1]，金帆[2]，冯德华[1]，王云龙[1]，梁毅军[1]*

1. 西安交通大学，西安 710049；2. 浙江科技学院，浙江 310023

摘　要：太阳黑子是发生在太阳光球层的现象，对太阳耀斑预测具有重要意义。针对类别样本数量不均衡的三分类太阳黑子数据集，提出了双模型集成分类算法。该方法通过一轻一重两个模型，分别承担两个类别的分类任务，再将二者的分类结果进行集成，夹逼出第三个类别的分类结果。实验表明该方法能够减少单个模型在不均衡数据集上出现的过拟合和欠拟合情况带来的不利影响，从新的角度解决了太阳黑子数据集的类别不均衡问题，平均F1分数达到0.931。

关键词：太阳黑子分类；双模型集成；类别不均衡；过拟合；欠拟合

中图法分类号：TP391.4

0　引　言

太阳黑子与太阳活动密切相关，太阳活动中最强烈的活动现象是太阳耀斑[1,2]，它主要爆发在太阳黑子上方的大气中。这些活动会扰乱地球大气层，影响地面无线电短波通信，并产生一些危害，如“磁暴”现象。世界各地的天文台一直在追踪所有可见的太阳黑子群，并对其进行分类和分析，以便及早发现耀斑。如果能够比较准确地实现太阳黑子群自动分类，则可以更好地检测某些类别太阳黑子群的生成，对预警耀斑具有重要意义，同时提高对太阳周期、空间气候及其对地球气候系统影响的理解和预警能力。典型的黑子群分类方法可以分为基于传统数字图像处理算法、数学形态法和小波分析法，以及当下流行的基于数据驱动的机器学习方法。近年来随着观测条件的进步，与太阳活动有关的数据增长迅速，基于深度学习方法的优势越来越显著。

Colak 等[3]采用 McIntosh 分类方案，提出了一种混合系统，该系统使用从 SOHO/MDI 磁图图像中提取的活动区域数据，对 SOHO/MDI 白光图像上的黑子群进行自动检测。从 MDI 白光图像中检测到黑子后，使用 MDI 磁图图像对它们进行分组/聚类。通过集成图像处理和神经网络，自动对检测到的黑子进行分类。但系统存在分组错误和小型黑子漏检的缺陷。Colak 等[4]使用机器学习方法结合传统图像处理算法提取太阳黑子特征，改善了特征提取效果，但仍存在在不同数据集上泛化性较差的问题。

Abd 等[5]采用改良后的七类 Zurich 分类方案，使用 SVM(Support Vector Machine，支持向量机)来实现对太阳全日面白光图像上黑子群的自动分类。在数据预处理阶段，使用了边缘检测、噪声去除以及二值化来分割黑子群和日面，再对黑子群进行无监督分割，将属于同一组的黑子进行合并，然后对每个黑子组的属性进行提取，最后使用 SVM 来进行分类。该方法精度对图片质量和失真程度有所要求，且分割过程对推断耗时有较大影响。

随着深度学习的发展和可训练数据量的增加，越来越多的人尝试基于卷积神经网络提取图像特征，以解决图像的分类[6,7]、分割[8,9]和检测[10,11]问题。Fang 等[12]使用 CNN(Convolutional

* 基金项目：国防科技创新特区 163 计划纵向研究课题，编号 19-163-21-TS-001-057-08

作者简介：陈清源，男，硕士。研究方向：嵌入式视觉系统、目标检测跟踪。Email:1463980813@qq.com

通信作者：梁毅军，男，博士。研究方向：机器视觉、计算机图形学。Email:liang.yj@mail.xjtu.edu.cn

Neural Network，卷积神经网络)对黑子群的磁类型进行分类，在数据预处理阶段图像被分为

* 基金项目：国防科技创新特区 163 计划纵向研究课题，编号 19-163-21-TS-001-057-08

作者简介：陈清源，男，硕士。研究方向：嵌入式视觉系统、目标检测跟踪。Email:1463980813@qq.com
通信作者：梁毅军，男，博士。研究方向：机器视觉、计算机图形学。Email:liang.yj@mail.xjtu.edu.cn

三类，白光图、磁图以及白光图和磁图的合成图像。将三类图像分别作为输入源输入卷积神经网络，进行分类，结果显示单独使用白光图的分类效果最佳。该方法作者认为磁图结构相对白光较为复杂，且磁图使用 CNN 不能很好地提取特征，是造成使用磁图分类不佳的主要原因之一；除此之外，三类图像样本数量不均衡带来的问题同样没有得到解决，存在 Beta-x 类别过拟合的现象。

Yang 等[13]提出了一种双流 CNN 的太阳黑子磁类型分类方法。该模型通过双流结构解决了白光图和磁图的多元输入问题，在 Alpha 类别上的分类准确率较高，但模型参数量较大，在 Beta-x 类别上依然存在过拟合问题。除了分类问题，当下最优的太阳黑子的检测模型同样存在少样本类别过拟合的问题，如 Fu 等[14]提出的 SunspotsNet 太阳黑子检测模型依然无法解决类别不均衡问题。

而本文提出的双模型集成算法则是通过两个模型分别承担 Alpha 类和 Beta-x 类的分类任务，即针对 Alpha 类别设计最为适合的分类模型和训练策略，针对 Beta-x 类别设计最为合适的分类模型和训练策略，并辅以大量针对不均衡数据集训练的调优技巧，最终对两个模型的结果进行集成，从而能够最大程度地解决数据集中类别不均衡的问题。上述双模型集成算法同当下主流的集成算法不同[15,16,17]，其中涉及到的两个模型均有明确的任务分工，集成单元更少、效果更优。另说明，本文提出的基于双模型集成的不平衡数据集定向分类模型取得了阿里天池太阳风暴识别和预警人工智能挑战赛的第二名，挑战赛地址：https://tianchi.aliyun.com/competition/entrance/531803/rankingList ；代码也已开源：https://github.com/qingyuanchen1997/Dual-Model-Integration。

1　本文方法

1.1 数据集介绍与分析

深度学习由数据驱动，故首先对本实验所用数据集进行介绍与分析，进而引出后续算法。

本实验所用数据集为 SOLAR-STORM1，由空间环境人工智能预警创新工坊整理提供，可在天池实验室(https://tianchi.aliyun.com/dataset/)公开下载。该数据集基于 Mount Wilson 黑子群磁类型分类方案，该方案根据磁场极性对黑子群进行分类。数据集共包含 Alpha、Beta 和 Beta-x 三个类别的太阳黑子数据样本，分别代表单极黑子群、极性之间具有简单且明确划分的双极黑子群和极性之间无明确划分的复杂双极黑子群，三个类别总计 14469 个数据样本。其中每个数据样本均包含两张图像：磁图和白光图，二者为一一对应关系。

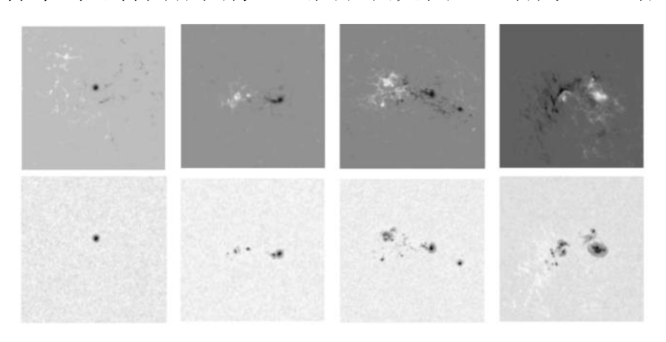


图 1 上图分别为 Alpha（左 1）、Beta（左 2）和更复杂的 Beta-x（左 3、左 4）磁类型的太阳黑子群磁场观测图像（行 1）和白光观测图像（行 2）

其中 Alpha 类太阳黑子的磁场图和白光图各有 4709 张，占比 32.54%，Beta 类太阳黑子的磁场图和白光图各有 7353 张，占比 50.82%，Beta-x 类太阳黑子的磁场图和白光图各有 2407 张，占比 16.64%，可见数据虽整体数量尚可，但三个类别的训练图像数量的差异却十分巨大，其中 Beta-x 类图像的数量远小于 Alpha 和 Beta 两个类别；另经查阅资料得知，在现实观测过程中，Alpha 类别和 Beta 类别的太阳黑子出现的概率也显著高于 Beta 类别的太阳黑

子出现的概率。

Fig 1 The above figure shows the magnetic field observation images of sunspot groups of Alpha (left 1), Beta (left 2) and more complex Beta-x (left 3 and left 4) magnetic types (line 1) and white light observation images (line 2)

表 1 Alpha、Beta、Beta-x 黑子数据样本占训练集比重

Table 1 Proportion of Alpha, Beta and Beta-x sunspot data samples in the training set

黑子类别	样本数量	数据集占比
Alpha 类	4709	32.54%
Beta 类	7353	50.82%
Beta-x 类	2407	16.64%

1.2 双模型集成算法

由于数据集在万张量级，总计 14469 组数据样本，为保证模型在验证集和测试集上取得结果的可靠性，故将其按 3:1:1 的比例划分为训练集、验证集和测试集。另由于数据具有时间连续性，其时间精度为 96 分钟，即每间隔 96 分钟对黑子群进行一次观测，且黑子群演化具有缓慢性，故同一个黑子群在一定时间内常会保持同一种磁类型，且相似度较高。为避免训练集同验证集、测试集存在信息上的重叠，实验基于时间段对数据集进行划分，将位于同一时间段内的不同样本统一放入训练集或验证集或测试集。文中所有实验均采用五折交叉验证，即将数据集切分为 5 份，1 份作为验证集，1 份作为测试集，剩余 3 份作为训练集，每组实验共进行 5 次，每次实验均基于已切分的 5 份数据集子集重构训练集、验证集和测试集，最终结果取 5 次实验的平均值。为保证双模型集成的可操作性，在每一轮训练中两个模型使用相同的数据集。实验首先使用 backbone 为 ResNet50 卷积层的网络在训练集上进行训练，并用训练出的模型对训练集和验证集进行分类。

表 2 使用 ResNet50 卷积层作为 backbone，Alpha 类、Beta 类、Beta-x 类在训练集和验证集上的 F1-score

Table 2 The F1-score of Alpha, Beta and Beta-x class on the training set and verification set while use ResNet50 as the backbone

黑子类别	训练集 F1-score	验证集 F1-score
Alpha 类	0.961	0.926
Beta 类	0.943	0.859
Beta-x 类	0.909	0.613

由表 2 可见，Beta-x 类在训练集上的 F1-score 远高于在验证集上的 F1-score，而 Alpha 类和 Beta 类在训练集上的 F1-score 只是略高于验证集上的 F1-score，由此可见基于 ResNet50 训练出的模型对于训练图片数量较少的 Beta-x 类存在过拟合的情况。

故实验采用相比于 ResNet50 卷积层数更少的 AlexNet 卷积层作为 backbone，从而得到更轻量级的网络模型，并重新进行了训练和分类。

表 3 使用 AlexNet 作为 backbone，Alpha 类、Beta 类、Beta-x 类在训练集和验证集上的 F1-score

Table 3 The F1-score of Alpha, Beta and Beta-x class on the training set and verification set while use AlexNet as the backbone

黑子类别	训练集 F1-score	验证集 F1-score
Alpha 类	0.934	0.868
Beta 类	0.899	0.821
Beta-x 类	0.876	0.753

结果如表 3 所示，更小更轻的网络很好地解决了 Beta-x 类因数量较少造成的过拟合问题，且显著提高了 Beta-x 类在验证集上的 F1-score。但由于此网络模型参数量较少，故模型的拟合能力较弱，导致训练图像数量较多的 Alpha 类和数量更多的 Beta 类出现了欠拟合的现象，其 F1-score 低于 ResNet50 模型分类得到的 F1-score。

由此可见，在分类数据集各类别数量极其不平衡的情况下，无法找出单一的一种神经网络模型同时兼顾训练图像数量较多的类别和训练图像数量较少的类别。即对于参数量较大的网络，其拟合能力较强，但需要较多的训练图像，此类模型对分类问题中训练图像数量较多的类别更有利；而对于参数量较少的网络，其拟合能力较弱，但需要较少的训练图像，此类模型对分类问题中训练图像数量较少的类别更有利；由此，本文确立了双模型集成理论上的有效性。

表 4 各网络模型对验证集的分类结果中 Alpha 类和 Beta-x 类相互被误分的图像数量

Table 4 The classification results of each network model on the verification set Alpha Class and Beta-Number of x-class images misclassified with each other

Backbone	Alpha 误分为 Beta-x case 数量	Beta-x 误分为 Alpha case 数量
ResNet18	0	1
ResNet50	0	0
AlexNet	0	0

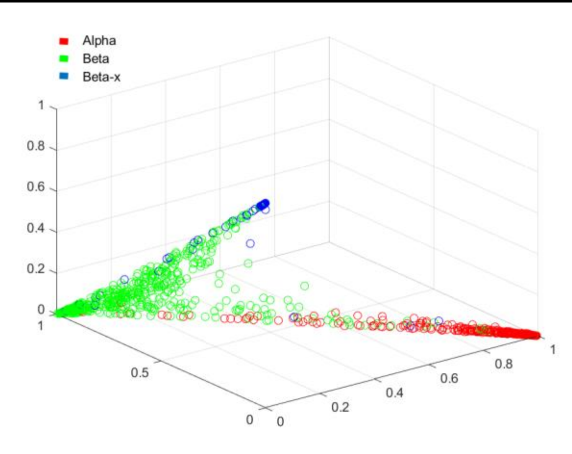


图 2 Alpha 类（红）、Beta 类（绿）、Beta-x 类（蓝）特征空间分布示意图

Fig 2 Schematic diagram of characteristic spatial distribution of Alpha(red), Beta(green) and Beta-x (blue)

另经过多个模型的测试，发现 Alpha 类和 Beta-x 类存在互不相干性；由表 4 可得，在所有模型的分类结果中，真实的 Alpha 类没有出现或极少出现被误分为 Beta-x 类的情况，同样，真实的 Beta-x 类没有出现或极少出现被误分为 Alpha 类的情况，两者的错误均是因为同 Beta 类的混淆。另外将利用神经网络提取出三维特征在三维特征空间中进行显示，由图 2 可以看出，Alpha 和 Beta-x 的样本点在特征空间中距离普遍较远，二者特征点极少发生混叠。故根据上述实验结果，确定了双模型集成的可行性，即不同模型的 Alpha 与 Beta-x 的结果互不影响，集成后的结果可以保留两个模型分别在 Alpha 和 Beta-x 上的分类精度；由此，本文确立了双模型集成的理论可行性。

综合上述情况，分别单独设计一重一轻两个模型；利用重模型的强拟合能力，令其专注于数量较多的 Alpha 类太阳黑子的分类；利用轻模型的抗过拟合能力，令其专注于数量较少的 Beta-x 类太阳黑子分类，最后再利用 Alpha 类和 Beta-x 类的互不相干性，对两个模型的分类结果进行不损失精度的集成融合，用重模型分出的 Alpha 类和轻模型分出的 Beta-x 类夹逼出 Beta 类，即将未被标记类别的样本划归为 Beta 类，如图三所示。



图 3 双模型集成示意图

Fig 3 Schematic diagram of dual model integration

基于 ResNet18 设计了双通道并行网络 paraResNet，将其作为参数较多的大模型去针对训练图像数量较多的 Alpha 类；另基于 AlexNet 设计了 miniAlexNet 模型，将其作为参数较少的小模型去针对训练图像数量较少的 Beta-x 类。

1.3 Alpha-model: paraResNet

1.3.1 网络结构

在确立了双模型集成策略后，首先针对性设计承担 Alpha 类别分类任务的模型。

比赛主办方提供的数据集中每个太阳黑子群类别均包含两种图像，磁场观测图像和和白光观测图像。

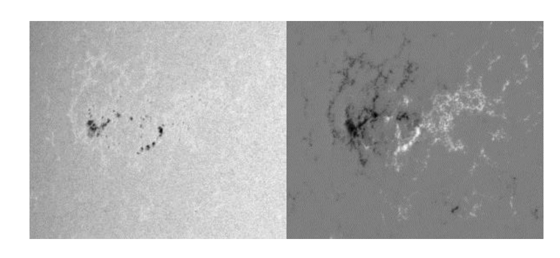


图 4 白光观测图像（左）和磁场观测图像（右）

Fig 4 The white light observation image(left) and the magnetic field observation image(right)

两图像尺寸相同，位深相同，故首先尝试将均为八位位深的单通道白光观测图像和磁场观测图像拼接成 16 位深图像送入 ResNet18 网络进行训练和分类，但是由于其纹理特征差异较大，拼接而成的双通道图像并不是一幅自然图像，其分类效果并不理想。故根据此数据集拥有白光、磁场两种图像的特点，为充分提取白光图和磁场图的特征，采用了同 Yang 等[13]提出的双流模型方案类似的方案，设计了双通道 ResNet18 并行网络以侧重于在 Alpha 类别上的分类性能；该网络的上下两个通道分别负责提取白光图和磁图的图像特征，而后将提取到的特征进行拼接后送入全连接网络。

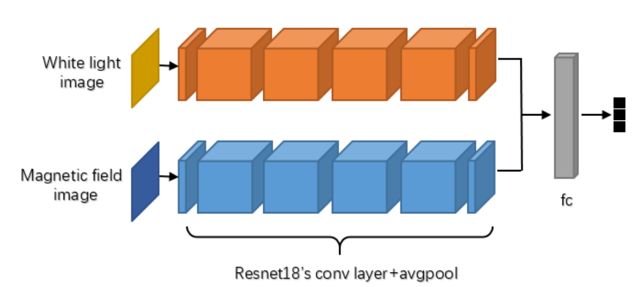


图 5 paraResNet 网络结构图

Fig 5 Network structure diagram of paraResNet

在此基础上，本实验参考了 He 等[18]在模型优化方向的工作，对 ResNet18 的下采样层进行改进，将原本同时兼顾减小特征图尺寸和改变特征张量深度的卷积层，替换成了负责减小特征图尺寸的平均池化层和负责改变特征张量深度的步长为 1 的卷积层，通过任务分工减轻了原卷积核的负担，提高了下采样模块的精度。根据该网络结构，本文将之称为 Parallel-ResNet18-D，缩写为 paraResNet。

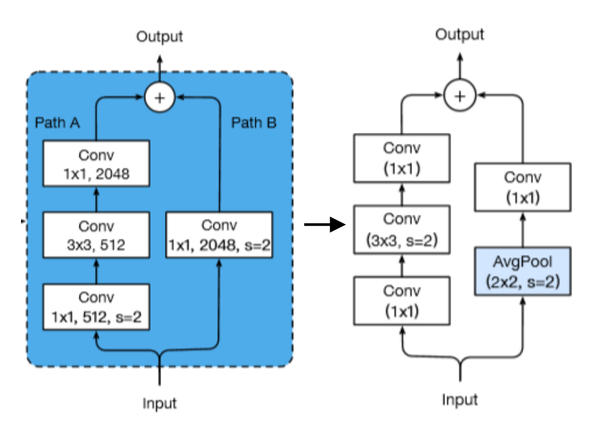


图 6 ResNet18 的下采样改进方案：RseNet-D

Fig 6 Improved down sampling scheme of ResNet18: RseNet-D

1.3.2 训练策略

在训练技巧上，首先使用单通道白光图和磁场图分别对两个单网络 ResNet18 模型进行预训练（不采用 ImageNet 预训练），再将两个预训练模型的卷积层迁移至双网络模型的并行卷积层。

在损失上，使用加权的交叉熵损失，样本较少的类别享有更大的权重，用以减小各类别样本数量不均衡带来的影响。

在数据增强方面，由于 Alpha 类和与其相邻的 Beta 类可用于训练的图像数量较多，故只采用了对图像信息损失较小的水平镜像和竖直镜像的数据增强方法；若再使用其他数据增强方法不仅无法提升模型分类正确率，反而会影响训练时间。

1.4 Beta-x-model: miniAlexNet

1.4.1 网络结构

针对训练样本较少的 Beta-x 类别，为减少参数量，未使用双通道并行网络结构，而是使用相比于 ResNet18 更为轻量级的 AlexNet[19]的卷积层作为 Beta-x-model 的 backbone，输入源为白光图和磁图的混合图像，即图像的第一个 Channle 为白光图，第二个 Channel 为磁图，为使用 ImageNet 数据集的预训练模型，需保证输入图像为常规三通道图像，故设置第三 Channel 为白光图和磁图各按 0.5 权重的加和，以引导网络学习白光图和磁图之间的纹理关联信息并填补通道空缺。为进一步减少后续全连接的参数量，在最后一个卷积层后接一个 output_size 为 3x3 的自适应池化层，最后将特征送入单层全连接，输出三个类别的置信度。同时，为了补偿网络提取图像特征的能力，为网络输入较高分辨率的图像，将原图像分辨率变换为 500×375。

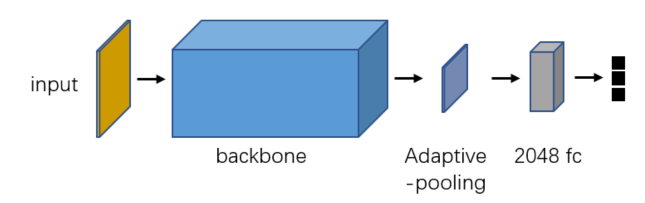


图 7 miniAlexNet 网络结构图

Fig 7 Network structure diagram of miniAlexNet

1.4.2 训练策略

首先，对数据集进行清洗。对于 Alpha 类别和 Beta 类别的数据，鉴于其时间轴上连续性较差，全部保留。对于 Beta-x 类别的数据，首先按照时间进行排序，根据该类别数据集图像序列在时间轴上的强连续性，对 Beta-x 类别的数据以 3/4 的下采样率进行数据清洗，大大降低了数据的冗余度和网络对重复信息进行无用学习的压力。

由于 Beta-x 类别和 Beta 类别之间的差异较小，故采用降温技巧，即降低原先输出的 logit 值，再进行 softmax，以增加网络在学习时各个类别的差异性，降低网络学习压力，加快收敛速度。

在数据增强上，采用了对图像信息影响较大但抗过拟合效果更佳的图像随机角度旋转对训练集进行增广，旋转角度控制在 90 度以内。

最后，通过对验证集的 logit 输出分布的分析，发现当 Beta-x 输出头的 logit 值同 Beta 输出头的 logit 值的差值在[0, 0.5] 区间内时，存在 7%的 Beta 类别样本被误判为 Beta-x 类别。故当二者所代表的输出头的 logit 值的差值在[0, 0.5]区间内时，增大 Beta 输出头的 logit 值，进行概率性校正。

1.5 模型集成

基于 1.3 节讨论的 Alpha 类和 Beta-x 类的不相干性，将 paraResNet 模型得到的分类结果同 miniAlexNet 模型的分类结果融合，即采用前者的 Alpha 类别分类结果和后者的 Beta-x 类别的分类结果，将未被标记类别的样本划归为 Beta 类，从而夹逼出 Beta 类别的分类结果。

2 实验与结果分析

2.1 参数设置

在针对 Alpha 的模型训练过程中，本实验采用了 Adam 自适应优化器[20]，初始学习率定为 2e-6；作为 SGD 的扩展，Adam Optimizer 能够根据每个参数的先前梯度，自适应地调整梯度的系数，适用于数据量较大的情况。损失上则使用加权的交叉熵损失，所加权重为 Alpha：1.56、Beta：1.0、Beta-x：3.05，该权重是基于 Alpha、Beta、Beta-x 三个类别在训练集中的样本数量比例进行设置，在该比例的基础上根据验证集分类效果进行微调，从而能够在损失层面降低训练数据类别不均衡带来的影响。

在针对 Beta-x 的模型训练过程中，本实验采用了带有动量的 SGD 下降策略，学习率为 0.0008，动量参数为 0.9。另外，本实验做了两个额外的数据集预处理操作，首先因 Beta-x 类别样本数据量本身较小，故将输入图像分辨率修改为 500×375，尽可能保留数据集信息；其次对训练集中额外的冗余信息进行滤除，即对 Beta-x 类别中在时间轴上连续的图像样本序列进行以 4 为步长的下采样，最终保留 891 个训练样本。

2.2 评价指标

本实验采用 F1-score 作为太阳黑子三分类任务的衡量指标，它是精确率和召回率的调和平均数，最大为 1，最小为 0；F1-score 接近 1 代表模型在该类别上的分类性能较好；反之，接近 0 则代表模型在该类别上的分类性能较差；F1-score 计算公式见式(1)。

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precission + recall} \#(1)$$

数据集共有三个类别 Alpha、Beta、Beta-x，每个类别均对应一个 F1-score，因为 Beta 类别太阳黑子在现实中出现概率最高，故本实验将 Beta 类别的 F1-score 作为首要考虑对象，其次为 Alpha 和 Beta-x。

2.3 消融实验

本论文针对训练数据不平衡的特性，采用了针对性的双模型融合算法，即使用双通道并行 ResNet18 进行针对性训练得到较好的 Alpha 类别分类模型，使用 miniAlexNet 进行针对性训练得到较好的 Beta-x 类别分类模型。最后再基于 Alpha 类别和 Beta-x 类别的不相干性，将前者的 Alpha 类别分类结果同后者的 Beta 类别分类结果进行融合，得到最终的三分类结果。

本论文对分类模型的评价指标同比赛方要求相同，使用综合了精确率和召回率的 F1-score 以衡量模型在测试集上的分类性能。

2.3.1 Alpha-model

首先对针对 Alpha 类别的模型进行消融实验，以验证各项策略的有效性。

表 5 各项技巧为 Alpha-model 带来的 F1-score 提升

Table 5 Improvement of F1-score brought by various skills to Alpha-model

Backbone: ResNet18				
数据增强（镜像）	✓	✓	✓	✓
双通道		✓	✓	✓
损失加权			✓	✓
下采样模块改进				✓
Alpha F1-score	0.945	0.969	0.967	**0.970**
Beta F1-score	0.849	0.867	0.883	**0.887**
Beta-x F1-score	0.685	0.676	0.711	**0.713**

注：加粗字体为每行最优值

由表 5 可见，每一项技巧都对模型的分类性能有了显著提高；数据增强（水平镜像+竖直镜像）在对图像信息影响较小的基础上弥补了训练集较小的问题，增强了抗过拟合能力；双通道网络更好地提取了太阳黑子的白光图像和磁场图像的特征，更多的参数也带来了更强的拟合能力；损失加权平衡了 Alpha 类别同 Beta 类别的在训练数据量上的不平衡关系；下采样模块在精度上的提高进一步提升了网络提取特征的有效性；最为关键的是，针对性训练后的双通道并行模型 paraResNet 对 Alpha 类别的分类效果非常好。

2.3.2 Beta-x-model

对针对 Beta-x 类别的模型进行消融实验，以验证各项策略的有效性。

表 6 各项技巧为 Beta-x-model 带来的 F1-score 提升

Table 5 Improvement of F1-score brought by various skills to Beta-x-model

Backbone: AlexNet				
数据增强(翻转+清洗)	✓	✓	✓	✓
高分辨率训练		✓	✓	✓
降温处理			✓	✓
概率性校正				✓
Alpha F1-score	0.872	0.907	0.927	**0.932**
Beta F1-score	0.839	0.872	0.891	**0.898**
Beta-x F1-score	0.786	0.814	0.853	**0.877**

注：加粗字体为每行最优值

由表 6 可见，每一项技巧都对模型的分类性能有了显著提高；数据增强（随机角度旋转）大大减轻了 Beta-x 类别数据量过小带来的不利影响，降低了小数据量带来的过拟合风险；将更高分辨率的图像送进网络训练，弥补了小网络的特征提取能力的不足；降温处理和概率性校正则通过人为调整干预，提升了模型对于复杂分类边界的拟合能力；而小巧的网络，较少的参数，使 miniAlexNet 模型对训练数据量较少的 Beta-x 类别的分类效果相比于双通道网络更好。

2.3.3 双模型集成

最终再将 paraResNet 和 miniAlexNet 两个模型进行集成，即将前者的 Alpha 类别分类结果同后者的 Beta-x 类别分类结果进行融合，将未被标记类别的样本划归为 Beta 类别，得到最终的三分类结果。

表 7 双模型集成在测试集上带来的 F1-score 提升

Table 7 Improvement of F1-score brought by dual model integration on verification set

	paraResNet	miniAlexNet	双模型融合
Alpha F1-score	**0.970**	0.932	**0.970**
Beta F1-score	0.887	0.898	**0.946**
Beta-x F1-score	0.713	**0.877**	**0.877**

注：加粗字体为每行最优值

由表 7 可以看出，融合后的模型完全保留了双通道并行模型 paraResNet 在 Alpha 类别上得到的 F1-score 和 miniAlexNet 模型在 Beta-x 类别上得到 F1-score，从而夹逼出比二者都更高的 Beta 类别的 F1-score，最终在测试集上，Alpha 类、Beta 类和 Beta-x 类的 F1-score 分别为 0.970、0.946、0.877。

2.4 与其它算法比较

本实验对 Fang 等[12]和 Yang 等[13]设计的网络结构以及当下主流网络结构 ResNet[21]在使用了翻转增强和损失加权的基础上，在 SOLAR-STORM1 数据集上进行实验验证。

表 8 多种模型实验结果

Table 8 Experimental results of various models

	Alpha F1-score	Beta F1-score	Beta-x F1-score	平均 F1-score
Fang's CNN	0.903	0.854	0.726	0.828
ResNet18	0.941	0.860	0.711	0.837
ResNet50	0.957	0.868	0.680	0.835
双流 ResNet50	0.966	0.879	0.706	0.850
paraResNet	**0.970**	0.887	0.713	0.857
miniAlexNet	0.932	0.898	**0.877**	0.902
双模型集成	**0.970**	**0.946**	**0.877**	**0.931**

注：加粗字体为每列最优值；所有参考模型均采用了镜像数据增强和损失加权的训练技巧

对比实验结果如表 8 所示，从表中可以看出，本文提出的 paraResNet 模型在 Alpha 类别上取得的效果非常优异，F1-score 达到了 0.970；同样的，本文提出的 miniAlexNet 网络在 Beta-x 类别上取得了优异的性能，F1-score 达到了 0.877；将 paraResNet 模型和 miniAlexNet 模型集成后，保留了前者在 Alpha 类别上的性能和后者在 Beta-x 上的性能，取得了最优的 Beta 类别分类结果，Beta 的 F1-score 达到了 0.946，远超所有单网络模型。

3 结 论

论文对太阳黑子数据集进行了分析，在论证了双模型集成有效性和可行性的基础上，通过分别训练针对 Alpha 类别的双通道并行模型 paraResNet 和针对 Beta-x 类别的 miniAlexNet 模型并将二者结果融合夹逼 Beta 类别的方法，巧妙地避免了重模型在少样本类别上的过拟合问题和轻模型在多样本类别上拟合能力不足的问题，且保留了二者在各自优势类别上的优异性能，从全新的角度降低了类别不均衡问题带来的影响，最终在 SOLAR-STORM1 数据集的三个类别上取得的平均 F1-score 超越了所有单网络模型。

双模型集成算法不仅对太阳黑子分类任务有效，其余任何三分类任务均可参考此集成方法，此算法对所有基于机器学习的、各类别样本数量不均衡的分类任务具有启发性意义。

为了进一步减少模型的复杂度、参数量和前向推理耗时，并降低模型的部署难度，未来可以考虑通过知识蒸馏和小样本学习的方式，使得单网络模型能够同时兼顾在多样本类别和少样本类别上的分类性能，从而替代双模型结构。

Classification of sunspot magnetic types based on dual model integration

Chen Qingyuan[1], Jin Fan[2], Feng Dehua[1], Wang Yunlong[1], Liang Yijun[1 *]

1. *Xi'an Jiaotong University, Xi'an 710049, China*
2. *Zhejiang University of Science and Technology, Zhejiang 310023, China*

Abstract: Sunspots occur in the solar photosphere and can make the prediction of solar flares. Aiming at the three classification sunspot data set with unbalanced number of category samples, a dual model integrated algorithm is proposed. This method uses two models, one light and one heavy, to undertake the classification tasks of two categories respectively, and then integrate the classification results of the two to squeeze out the classification results of the third category. Experiments show that this method can reduce the adverse effects of over-fitting and under-fitting of a single model on unbalanced data sets, and solve the problem of class imbalance in sunspot data sets from a new perspective, with an average F1 score of 0.931.

Key words: Sunspot classification; Dual model integration; Category imbalance; Overfitting; Underfitting

参考文献(References)

[1] 梁波,林语琦,戴伟,冯松,杨云飞. 基于多变量 LSTM 网络的太阳黑子活动预测分析[J]. 天文研究与技术, 2020, 17(03):322-333.

Liang Bo, Lin Yuqi, Dai Wei, et al. Prediction and Analysis of Sunspot Activity Based on Multivariable LSTM Network[J]. Astronomical Research & Technology, 2020, 17(03):322-333.

[2]朱健,杨云飞,苏江涛,刘海燕,李小洁,梁波,冯松. 基于深度学习的太阳活动区检测与跟踪方法研究[J]. 天文研究与技术, 2020, 17(02):191-200.

Zhu Jian, Yang Yunfei, Su Jiangtao, et al. A Detection and Tracking Method for Active Regions Based on Deep Learning[J]. Astronomical Research & Technology, 2020, 17(02):191-200.

[3] Colak T and Qahwaji R. Automated mcintosh-based classification of sunspot groups using MDI images[J]. Solar Physics, 2008, vol.248, no.2, pp.277–296.

[4] Colak T and Qahwaji R. Automated solar activity prediction: a hybrid computer platform using machine learning and solar imaging for automated prediction of solarflares[J]. Space Weather, 2009, 7(6).

[5] Abd M A, Majed S F, and Zharkova V. Automated classification of sunspot groups with support vector machines[J]. Proceedings of the Technological Developments in Networking, Education and Automation, 2010, pp.321–325.

[6]Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, and Rabinovich A. Going deeper with convolutions[C]// Computer Vision and Pattern Recognition. 2015, Boston, United States: IEEE: 1-9.

[7]张珂,冯晓晗,郭玉荣,苏昱坤,赵凯,赵振兵,马占宇,丁巧林. 图像分类的深度卷积神经网络模型综述[J]. 中国图象图形学报，2021, 26(10):2305-2325.

Zhang Ke, Feng Xiaohan, Guo Yurong, Su Yukai, Zhao Kai, Zhao Zhenbing, Ma Zhanyu and Ding Qiaolin. Overview of deep convolutional neural networks for image classification[J]. Journal of Image and Graphics, 2021, 26(10):2305-2325

[8] Ronneberger O. U-Net Convolutional Networks for Biomedical Image Segmentation[C]// Bildverarbeitung für die Medizin 2017. 2017, Berlin, Germany: Springer Berlin Heidelberg: 3-3.
[9]周涛,董雅丽,霍兵强,刘珊,马宗军. U-Net 网络医学图像分割应用综述[J]. 中国图象图形学报, 2021, 26(09):2058-2077.
Zhou Tao, Dong Yali, Huo Bingqiang, Liu Shan and Ma Zongjun. U-Net and its applications in medical image segmentation: a review[J]. Journal of Image and Graphics, 2021, 26(09) :2058- 2077.
[10] Cai Z W and Vasconcelos N. Cascade R-CNN: Delving Into High Quality Object Detection[C]// Computer Vision and Pattern Recognition. 2018, Salt Lake City, United States: IEEE: 1483-1498.
[11]赵永强,饶元,董世鹏,张君毅. 深度学习目标检测方法综述[J]. 中国图象图形学报, 2020, 25(04):0629-0654.
Zhao Yongqiang, Rao Yuan, Dong Shipeng P and Zhang Junyi. Survey on deep learning object detection[J]. Journal of Image and Graphics, 2020, 25(04): 0629-0654.
[12] Fang Y H, Cui Y M and Ao X Z. Deep Learning for Automatic Recognition of Magnetic Type in Sunspot Groups[J]. Advances in Astronomy. 2019, 2019(123):1-10.
[13] 杨五谷,田卫新,曾曙光,黄瑶. 一种双流 CNN 的太阳黑子群磁类型分类方法[J]. 天文研究与技术，2021, https://doi.org/10.14005/j.cnki.issn1672-7673.20210823.003.
Yang Wugu, Tian Weixin, Zeng Shuguang and Huang Yao. A magnetic classification method of sunspot group based on dual-stream CNN[J]. Astronomical Research & Technology, 2021, https://doi.org/10.14005/j.cnki.issn1672-7673.20210823.003.
[14]付小娜廖成武,白先勇,梁波,冯松,杨洪娟,杨云飞. 基于 LeNet-5 卷积神经网络的太阳黑子检测方法[J]. 天文研究与技术, 2018, 1672－7673(2018) 03－0340－07.
Fu Xiaona, Liao Chengwu, Bai Xianyong, Liang Bo, Feng Song, Yang Hongjuan and Yang Yunfei. Sunspot detection method based on lenet-5 convolutional neural network[J]. Astronomical Research & Technology, 2018, 1672－7673(2018) 03－0340－07.
[15] Chen T Q and Carlos G. XGBoost: A Scalable Tree Boosting System[C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, San Francisco, United States: ACM: 785-794.
[16] Breiman L. Bagging predictors[J]. Machine Learning. 1996, 24: 123-140.
[17] Breiman L. Random Forest[J]. Machine Learning. 2001, 45(1): 5-32.
[18] He T, Zhang Z, Zhang H, Zhang Z Y, Xie J Y and Li M. Bag of Tricks for Image Classification with Convolutional Neural Networks[C]// Computer Vision and Pattern Recognition. 2019, Long Beach, United States: IEEE: 558-567.
[19] Krizhevsky A, Sutskever I, and Hinton G. ImageNet Classification with Deep Convolutional Neural Networks// Neural Information Processing Systems. 2017, New York, United States: ACM: 84-90.
[20] Kingma D P and Ba J. Adam: a method for stochastic optimization[EB/OL][R]. 2017, [2020-07-11]. https://arxiv.org/pdf/1412.6980.pdf.
[21] Zhang X Y, Ren S Q and Sun J. Deep Residual Learning for Image Recognition[C]// Computer Vision and Pattern Recognition. 2016, Las Vegas, United States: IEEE: 770-778.